

Specification

Image Comparison System and Image Comparison Method

Technical Field

5 The present invention relates to an image comparison system, image comparison method, and image comparison program and, more particularly, to an image comparison system, image comparison method, and image comparison program capable of comparison and search at
10 high accuracy even when no three-dimensional model of an object can be registered beforehand, only one or few reference images of each object exist in a database of the system or on a network, and these reference images are captured under different conditions such as pose and
15 illumination.

Background Art

Fig. 25 is a block diagram showing a conventional image comparison system. This conventional image comparison system comprises an image input means
20 115, image converting means 117, image comparing means 157, reference image storing unit 130, and standard three-dimensional object model storing unit 135.

Reference images obtained by sensing the images of objects are prestored in the reference image
25 storing unit 130. Standard three-dimensional object models are prestored in the standard three-dimensional object model storing unit 135. The image converting

means 117 generates a partial image by converting one or both of an input image from the image input means 115 and each reference image obtained from the reference image storing unit 130 such that the pose conditions are the same, in relation to a partial image common to the input image and reference image, by using a three-dimensional object model obtained from the standard three-dimensional object model storing unit 135.

10 A partial region is a characteristic portion such as an eye, nose, or mouth. Partial regions can be related to each other by designating feature points in images and three-dimensional object models beforehand. The image comparing means 157 compares the partial
15 images of the input image and each reference image converted by the image converting means 117, calculates average similarity degrees, and selects a reference image having the largest similarity degree for each object (e.g., Japanese Patent Laid-Open No. 2000-322577
20 (patent reference 1)).

Fig. 26 is a block diagram showing another conventional image comparison system. This conventional image comparison system comprises an image input means 115, illumination change correcting means 122, image
25 converting means 118, image comparing means 158, reference image storing unit 130, and standard three-dimensional object model storing unit 135.

Reference images obtained by sensing the images of objects are prestored in the reference image storing unit 130. Standard three-dimensional object models are prestored in the standard three-dimensional object model storing unit 135. The illumination change correcting means 122 estimates the illumination condition (surface reflectance) of an input image from the image input means 115, by using a three-dimensional object model obtained from the standard three-dimensional object model storing unit 135. The image converting means 118 generates an image by converting the input image by using a three-dimensional object model so as to meet the illumination condition of a reference image. The image comparing means 158 compares the input image converted by the image converting means 118 with individual reference images, calculates similarity degrees, and selects a reference image having the largest similarity degree for each object (e.g., Japanese Patent Laid-Open No. 2002-024830 (patent reference 2)).

Fig. 27 is a block diagram showing still another conventional image comparison system. This conventional image comparison system comprises an image input means 115, reference three-dimensional object model storing unit 137, and pose estimating/comparing means 150. The pose estimating/comparing means 150 includes a pose candidate deciding means 120, comparison

image generating means 140, and image comparing means 155.

Reference three-dimensional object models generated by measuring objects are prestored in the reference three-dimensional object model storing unit 137. The pose estimating/comparing means 150 calculates a minimum distance value (or a maximum similarity degree) between an input image obtained from the image input means 115 and each reference three-dimensional object model obtained from the reference three-dimensional object model storing unit 137, and selects a model having the smallest minimum distance value.

More specifically, the pose candidate deciding means 120 generates at least one pose candidate. Then, the comparison image generating means 140 generates a comparison image close to an input image while projecting a reference three-dimensional object model onto a two-dimensional image in accordance with the pose candidate. The image comparing means 155 calculates a distance value between the comparison image and the input image, and selects a comparison image having the smallest distance value for each model, thereby estimating an optimum pose and obtaining a minimum distance value between the input image and the reference three-dimensional object model. In addition, the image comparing means 155 selects a model having the smallest

minimum distance value (e.g., Japanese Patent Laid-Open No. 2003-058896 (patent reference 3)).

Disclosure of Invention

Problems to be Solved by the Invention

5 The image comparison systems of patent references 1 and 2 cannot perform comparison at sufficiently high accuracy if the poses or illumination conditions of an input image and reference image are different.

10 The reasons are as follows. In patent reference 1, image conversion is performed by matching the pose conditions by estimating the pose of each image. However, it is difficult to accurately estimate the pose of an image, so images cannot be correctly
15 matched. In addition, image conversion is performed using a standard three-dimensional object model different from the three-dimensional shape of an object to be observed. Therefore, distortion by image conversion increases if the shape is complicated or the
20 pose conditions are largely different.

 In patent reference 2, image conversion is performed by estimating the illumination conditions by using a standard three-dimensional object model different from the three-dimensional shape of an object
25 to be observed. Accordingly, a correction error sometimes occurs in a fine portion although the correction is roughly accurate.

Furthermore, the system of patent reference 3 cannot perform comparison if no three-dimensional object model of each object is registered beforehand or the number of reference images is small.

5 The reasons are as follows. In the technique of patent reference 3, three-dimensional object models are registered in advance and compared with an input image. Also, to register three-dimensional object models in advance, each object must be measured by a
10 three-dimensional shape measurement device before comparison. However, this measurement is difficult to perform in many cases. It is also possible to generate a three-dimensional object model from a plurality of images, but the generation of a three-dimensional object
15 model is difficult if the number of reference images is small.

Means for Solving the Problems

 The present invention has been made in consideration of the above situation, and has as its
20 object to make high-accuracy comparison and search possible even when reference images of objects are captured under different conditions such as pose and illumination.

 It is another object of the present invention
25 to make high-accuracy comparison and search possible even when no three-dimensional object model of each object can be obtained beforehand.

It is still another object of the present invention to make high-accuracy comparison and search possible even when each object has only one or few reference images.

5 To achieve the above objects, an image comparison system according to the present invention is characterized by comprising input means for input three-dimensional data of an object, reference image storing means for storing a reference image of at least
10 one object, pose candidate generating means for generating a pose candidate as a candidate for pose of the object, comparison image generating means for generating a comparison image close to the reference image while projecting the three-dimensional data onto a
15 two-dimensional image in accordance with the pose candidate, and image comparing means for performing comparison on the basis of one of a distance value and a similarity degree between the reference image and the comparison image.

20 An image comparison method according to the present invention is characterized by comprising the steps of input three-dimensional data of an object, generating a pose candidate as a candidate for pose of the object, generating a comparison image close to a
25 reference image while projecting the three-dimensional data onto a two-dimensional image in accordance with the pose candidate, and performing comparison on the basis

of one of a distance value and a similarity degree between the reference image and the comparison image.

An image comparison program according to the present invention is characterized by causing a computer to execute a procedure of input three-dimensional data of an object, a procedure of generating a pose candidate as a candidate for pose of the object, a procedure of generating a comparison image close to a reference image while projecting the three-dimensional data onto a two-dimensional image in accordance with the pose candidate, and a procedure of performing comparison on the basis of one of a distance value and a similarity degree between the reference image and the comparison image.

Effects of the Invention

The first effect of the present invention is that comparison and search can be performed at high accuracy even when reference images of each object are captured under different conditions such as pose and illumination. This is so because three-dimensional data of an object is measured, a comparison image matching the image sensing conditions such as the pose and illumination of each reference image is generated, and comparison is performed by comparing the comparison image with the reference image.

The second effect is that comparison and search can be performed at high accuracy even when no

three-dimensional object model of each object can be obtained beforehand, or even when only one or few reference images exist. This is so because three-dimensional data of an object is measured, a
5 comparison image matching the existing reference image is generated, and comparison is performed by comparing the comparison image with the reference image.

Brief Description of Drawings

Fig. 1 is a block diagram showing the
10 arrangement of the first embodiment of an image comparison system according to the present invention;

Fig. 2 is a block diagram showing the arrangement of an image comparing means of the first embodiment;

15 Fig. 3 is a flowchart showing the operation of 1 : 1 comparison in the first embodiment;

Fig. 4 is a flowchart showing the operation of 1 : N comparison in the first embodiment;

20 Fig. 5 is a view showing practical examples of reference images in the first embodiment;

Fig. 6 is a view showing a practical example of three-dimensional data in the first embodiment;

Fig. 7 is a view showing practical examples of comparison images in the first embodiment;

25 Fig. 8 is a block diagram showing the arrangement of the second embodiment of the present invention;

Fig. 9 is a flowchart showing the operation of
1 : N comparison in the second embodiment;

Fig. 10 is a block diagram showing the
arrangement of the third embodiment of the present
5 invention;

Fig. 11 is a block diagram showing the
arrangement of an image comparing means of the third
embodiment;

Fig. 12 is a flowchart showing the operation
10 of 1 : N comparison in the third embodiment;

Fig. 13 is a view showing practical examples
of standard three-dimensional reference points in the
third embodiment;

Fig. 14 is a view showing a practical example
15 of a standard three-dimensional weighting coefficient in
the third embodiment;

Fig. 15 is a view showing practical examples
of reference weighting coefficients in the third
embodiment;

Fig. 16 is a view showing practical examples
20 of input three-dimensional reference points in the third
embodiment;

Fig. 17 is a view showing practical examples
of two-dimensional weighting coefficients in the third
25 embodiment;

Fig. 18 is a block diagram showing the
arrangement of the fourth embodiment of the present

invention;

Fig. 19 is a flowchart showing the operation of the fourth embodiment;

Fig. 20 is a flowchart showing the operation
5 of the fourth embodiment;

Fig. 21 is a view showing practical examples of representative three-dimensional object models in the fourth embodiment;

Fig. 22 is a block diagram showing the
10 arrangement of the fifth embodiment of the present invention;

Fig. 23 is a flowchart showing the operation of the fifth embodiment;

Fig. 24 is a view showing practical examples
15 of representative images in the fifth embodiment;

Fig. 25 is a block diagram showing a conventional image comparison system;

Fig. 26 is a block diagram showing another conventional image comparison system; and

20 Fig. 27 is a block diagram showing still another conventional image comparison system.

Best Mode for Carrying Out the Invention

Embodiments of the present invention will be described in detail below with reference to the
25 accompanying drawings.

(First Embodiment)

Fig. 1 is a block diagram showing the first

embodiment of an image comparison system according to the present invention. In Fig. 1, reference numeral 10 denotes a three-dimensional data input means for input three-dimensional data; 30, a reference image storing unit; and 50, a pose estimating/comparing means. The pose estimating/comparing means 50 includes a pose candidate deciding means 20, comparison image generating means 40, and image comparing means 55.

Outlines of the operations of these means are as follows. First, a reference image obtained by sensing the image of at least one object is prestored in the reference image storing unit 30. The image sensing conditions such as the pose and illumination of the reference image are not limited. Also, the reference image storing unit 30 can be incorporated into the system or connected across a network to the system.

The three-dimensional data input means 10 inputs three-dimensional data of an object to be compared (or an object to be searched). Three-dimensional data can be generated by using, e.g., a three-dimensional shape measurement apparatus described in Japanese Patent Laid-Open No. 2001-12925, or an apparatus described in Japanese Patent Laid-Open No. 9-91436 which restores a three-dimensional shape from a plurality of images captured by a large number of cameras.

The pose estimating/comparing means 50

calculates a minimum distance value (or a maximum similarity degree) between the three-dimensional data input by the three-dimensional data input means 10 and the reference image obtained from the reference image storing unit 30. More specifically, the pose candidate deciding means 20 generates a pose candidate as a candidate for the pose of at least one object (the pose of an object is represented by the position and direction of the object). The comparison image generating means 40 generates a comparison image close to the reference image while projecting the three-dimensional data onto a two-dimensional image in accordance with the pose candidate.

The image comparing means 55 includes a calculator 55a, selector 55b, and comparator 55c shown in Fig. 2. The image comparing means 55 calculates a distance value between the comparison image and the reference image in the calculator 55a, and selects a comparison image having the smallest distance value with respect to the reference image in the selector 55b, thereby estimating an optimum pose and obtaining a minimum distance value between the three-dimensional data and the reference image. In addition, in comparison (1 : 1 comparison) with one object (reference image), the comparator 55c compares the minimum distance value with a threshold value to determine whether the two objects are the same. That is, if the minimum

distance value is equal to or smaller than the threshold value, the comparator 55c determines that the two objects are the same; if not, the comparator 55c determines that the two objects are different. Also, in
5 a process (1 : N comparison) of searching a plurality of objects for an object (reference image) closest to the input three-dimensional data, the comparator 55c selects a reference image having the smallest minimum distance value. Note that when determination is to be performed
10 by using a similarity degree between a comparison image and a reference image, it is determined that the two objects are the same if the similarity degree is equal to or larger than a threshold value, and it is determined that the two objects are different if not.

15 The overall operation of 1 : 1 comparison of this embodiment will be explained in detail below with reference to Fig. 1 and a flowchart in Fig. 3. In this explanation, comparison of input three-dimensional data with a reference image R_k will be described.

20 Referring to Fig. 3, the three-dimensional data input means 10 inputs three-dimensional data (step 100). Then, the pose candidate deciding means 20 decides a pose candidate group $\{e_j\}$ (step 110). The comparison image generating means 40 generates a
25 comparison image close to the reference image R_k while projecting the three-dimensional data onto a two-dimensional image in accordance with a pose

candidate (step 120). The image comparing means 55 calculates a distance value between the comparison image and the reference image (step 130). Furthermore, the image comparing means 55 selects a comparison image
5 having the smallest distance value, thereby estimating an optimum pose and obtaining a minimum distance value between the three-dimensional value and the reference image R_k (step 140).

Although a pose candidate having the smallest
10 distance value is selected from the predetermined pose candidate group in the above description, it is also possible to return to the pose candidate deciding means 20 to search for a pose candidate having the smallest distance value while sequentially changing pose
15 candidates. Then, the pose estimating/comparing means 50 compares the minimum distance value with a threshold value to determine whether the two objects are the same (step 155).

The overall operation of 1 : N comparison of
20 this embodiment will be explained in detail below with reference to Fig. 1 and a flowchart in Fig. 4.

Referring to Fig. 4, the three-dimensional data input means 10 inputs three-dimensional data (step 100). Then, the pose estimating/comparing means 50 initially
25 sets the image number of a reference image to $k = 1$ (step 150). After that, the same processes as in steps 110 to 140 of Fig. 3 are executed in steps 110 to 140.

Subsequently, the pose estimating/comparing means 50 increments the image number k by 1 (step 151), and compares the image number k with an image count M (the number of reference images) (step 152). If the
5 image number k is equal to or smaller than the image count M , the flow returns to step 110 to perform the same processing, thereby calculating the minimum distance value of the next reference image. Finally, if the image number k is equal to or larger than the image
10 count M in step 152, a reference image R_k having the smallest minimum distance value is output as a comparison result (step 153).

The operation of this embodiment will be explained in more detail below by using a practical
15 example. Although an explanation will be made by taking a person's face as an example of an object of comparison, the present invention is of course also applicable to comparison of another object. First, as shown in Fig. 5, reference images $R_k(r)$ (r is the index
20 of a pixel or feature) of objects k are stored in the reference image storing unit 30. Note that the pose conditions of these images are not necessarily the same (differences between illumination conditions are not shown). Note also that one reference image is stored
25 for each object in this example, but a plurality of reference images may also be stored for each object.

The operation of 1 : N comparison will be

described below. First, assume that three dimensional data as shown in Fig. 6 is input from the three-dimensional data input means 10 (step 100 in Fig. 4). As shown in Fig. 6, this three-dimensional data has a shape $P_0(x,y,z)$ and texture $T_0(R,G,B)$ in a three-dimensional space (x,y,z) on the object surface as information. Q represents the index of a point on the object surface, and corresponds to, e.g., the coordinates of a point $Q(s,t)$ which is obtained by projecting a point on the object surface from the barycenter of the object onto a sphere having the barycenter of the object as its center. To increase the efficiency of comparison, CG images for learning are generated beforehand from three-dimensional data under various illumination conditions by computer graphics, and a base image group is obtained by analyzing the main components of these CG images for learning.

Then, the pose estimating/comparing means 50 initially sets the image number of a reference image to $k = 1$ (step 150). The pose candidate deciding means 20 decides a pose candidate group $\{e_j\}$ (step 110). Although the pose candidate group can be preset regardless of a reference image, it is also possible to manually or automatically extract reference points such as an eye, nose, and mouth from a reference image and three-dimensional data, and estimate a rough pose by using a method of calculating the position and direction

of an object described in Japanese Patent Laid-Open
No. 2001-283229, thereby generating a pose candidate
group around the estimated pose. A rough pose may also
be estimated and stored in advance by comparing
5 representative three-dimensional data (model) prepared
beforehand with a reference image, without using any
input three-dimensional data.

Then, the comparison image generating means 40
generates comparison images $G_{kj}(r)$ close to the
10 illumination condition of the reference image R_k while
projecting the three-dimensional data onto a
two-dimensional image in accordance with a pose
candidate e_j (step 120). These comparison images close
to the illumination condition of the reference image can
15 be generated by converting the coordinates of previously
obtained base images on the basis of each pose
candidate, and calculating the coefficient of the linear
sum of the base images having the converted coordinates
by the least square method such that the linear sum is
20 close to the input image. Fig. 7 shows examples of
comparison images generated for a reference image R_1 (no
grayscale information is shown).

Subsequently, the image comparing means 55
calculates a distance value between the comparison image
25 and the reference image (step 130). When a Euclid
distance is used, for example, the distance value can be
calculated by

$$D_{kj} = \sum_r \{R_k(r) - G_{kj}(r)\}^2$$

When a similarity degree S_{kj} is used, for example, the similarity degree S_{kj} can be calculated by

$$S_{kj} = \exp(-D_{kj})$$

5 In this case, the pose estimating/comparing means 50 estimates an optimum pose by selecting a comparison image having the smallest distance value, and also calculates a minimum distance value D_k between the three-dimensional data and the reference image R_k by

10 minimum distance value $D_k = \min_j D_{kj}$

(step 140). In the case shown in Fig. 7, G_{11} , for example, is selected.

Then, the image number k is incremented by 1 (step 151), and the image number k is compared with the
 15 image count M (step 152). If the image number k is equal to or smaller than the image count M , the flow returns to step 110 to perform the same processing, thereby calculating the minimum distance value of the next reference image. Finally, if the image number k is
 20 equal to or larger than the image count M in step 152, a reference image R_k having the smallest minimum distance value is output as a comparison result (step 153). In the case of the three-dimensional data shown in Fig. 6, if the minimum distance values for the reference images
 25 R_k are calculated as $\{20, 50, 25\}$, for example, the reference image R_1 shown in Fig. 5 is selected as a reference image having the smallest minimum distance

value.

Note that a distance value between a comparison image and a reference image is calculated in this embodiment, but a similarity degree may also be used instead of the distance value. As an example, the similarity degree can be calculated by the method as described above. When the similarity degree is used, a reference image having the largest maximum similarity degree is output as a comparison result.

10 This applies to all embodiments to be described below.

 In this embodiment, three-dimensional data of an object is measured, and this three-dimensional data is compared and compared with a reference image while the pose and illumination condition are corrected. Therefore, comparison and search can be performed at high accuracy even when reference images of objects are captured under different conditions such as pose and illumination. Also, since three-dimensional data of an object is measured and this three-dimensional data is compared and compared with a reference image, comparison and search can be performed at high accuracy even when no three-dimensional object model of each object can be obtained beforehand, or even when only one or few reference images exist.

25 (Second Embodiment)

 Fig. 8 is a block diagram showing the second

embodiment of the present invention. The same reference numerals as in Fig. 1 denote the same parts in Fig. 8. This embodiment comprises a three-dimensional data input means 10, reference image storing unit 30, pose
5 estimating/comparing means 51, and reference correction coefficient storing unit 65. The pose estimating/comparing means 51 includes a pose candidate deciding means 20, comparison image generating means 40, image comparing means 55, and score correcting means 60.
10 The difference from Fig. 1 is the addition of the score correcting means 60 and reference correction coefficient storing unit 65.

Outlines of the operations of these means are as follows. First, the three-dimensional data input
15 means 10, reference image storing unit 30, pose candidate deciding means 20, comparison image generating means 40, and image comparing means 55 perform the same processes as in the first embodiment shown in Fig. 1.

A coefficient for correcting a comparison
20 score (a distance value or similarity degree) in relation to a reference image is prestored in the reference correction coefficient storing unit 65. The pose estimating/comparing means 51 calculates a minimum distance value (or a maximum similarity degree) between
25 three-dimensional data input from the three-dimensional data input means 10 and a reference image obtained from the reference image storing unit 30, and also corrects

this minimum distance value by using the correction coefficient obtained from the reference correction coefficient storing unit 65.

More specifically, the pose candidate deciding means 20 generates at least one pose candidate. The comparison image generating means 40 generates a comparison image close to a reference image while projecting three-dimensional data onto a two-dimensional image in accordance with the pose candidate. The image comparing means 55 calculates a distance value between the comparison image and the reference image, and selects a comparison image having the smallest distance value for each reference image, thereby estimating an optimum pose and obtaining a minimum distance value between the three-dimensional data and the reference image.

The score correcting means 60 corrects the minimum distance value by using a correction coefficient corresponding to the reference image. Also, in comparison (1 : 1 comparison) with one object (reference image), the score correcting means 60 compares the corrected minimum distance with a threshold value to determine whether the two objects are the same. Furthermore, in a process (1 : N comparison) of searching a plurality of objects for an object (reference image) closest to input three-dimensional data, the score correcting means 60 selects a reference

image having the smallest corrected minimum distance value.

The overall operation of 1 : N comparison of this embodiment will be explained in detail below with reference to Fig. 8 and a flowchart in Fig. 9.

Referring to Fig. 9, the three-dimensional data input means 10 inputs three-dimensional data (step 100). Then, the pose estimating/comparing means 50 initially sets the image number of a reference image to $k = 1$ (step 150). Also, the pose candidate deciding means 20, comparison image generating means 40, and image comparing means 55 estimate an optimum pose and calculate a minimum distance value between the three-dimensional data and a reference image R_k by the same processes as in the first embodiment (steps 110 to 140).

Subsequently, the score correcting means 60 corrects the minimum distance value by using a correction coefficient corresponding to the reference image R_k (step 160). The pose estimating/comparing means 51 increments the image number k by 1 (step 151), and compares the image number k with an image count M (step 152). If the image number k is equal to or smaller than the image count M , the flow returns to step 110 to perform the same processing, thereby calculating the minimum distance value of the next reference image, and correcting the minimum distance value by using a

correction coefficient corresponding to the reference image. If the image number k is equal to or larger than the image count M in step 152, a reference image R_k having the smallest corrected minimum distance value is
5 output as a comparison result (step 153).

In 1 : 1 comparison, the processes in steps 100, 110, 120, 130, 140, and 160 are performed in the same manner as in the first embodiment, and the process in step 155 of Fig. 3 is performed after that. In step
10 155 of Fig. 3, whether the two objects are the same is determined by comparing the distance value with a threshold value as described earlier.

The operation of this embodiment will be explained in more detail below by using a practical
15 example. First, as in the explanation of the first embodiment, reference images $R_k(r)$ as shown in Fig. 5 are stored in the reference image storing unit 30. Also, correction coefficients as shown in Table 1 are stored in the reference correction coefficient storing
20 unit 65. For example, a correction coefficient A_k is obtained from $A_k = A/E_k$ (A is a preset constant) by calculating a minimum distance value D_k^h between each representative three-dimensional data (representative three-dimensional object model) C_h ($h = 1, \dots, H$)
25 prepared beforehand and a reference image R_k by the image comparison system of the first embodiment, and calculating average value $E_k = \sum_h D_k^h / H$ of the minimum

distance values (or the average value of high-order candidates having small minimum distance values). For example, if $A = 20$ and the average values E_k of the minimum distance values of individual reference images are {50, 40, 20}, the correction coefficients A_k are as shown in Table 1. Table 1 indicates that, for example, a reference image R_1 is captured under bad conditions and has a large average distance value.

[Table 1]

Image number	Correction coefficient
1	0.4
2	0.5
3	1.0

Assume that three-dimensional data as shown in Fig. 6 is input from the three-dimensional data input means 10 (step 100 in Fig. 9). The pose estimating/compared means 50 initially sets the image number of a reference image to $k = 1$ (step 150). Then, the pose candidate deciding means 20, comparison image generating means 40, and image comparing means 55 estimate an optimum pose and calculate a minimum distance value between the three-dimensional data and the reference image R_k by the same processes as in the first embodiment (steps 110 to 140).

Subsequently, the score correcting means 60 corrects the minimum distance value by using a correction coefficient corresponding to the reference

image R_k (step 160). For example, the minimum distance value D_k can be corrected by $D_k' = A_k D_k$ by using the correction coefficient A_k . Also, the pose estimating/comparing means 51 increments the image number k by 1 (step 151), and compares the image number k with the image count M (step 152). If the image number k is equal to or smaller than the image count M , the flow returns to step 110 to perform the same processing, thereby calculating the minimum distance value of the next reference image, and correcting the minimum distance value by using a correction coefficient corresponding to the reference image. Finally, if the image number k is equal to or larger than the image count M in step 152, a reference image R_k having the smallest corrected minimum distance value is output as a comparison result (step 153). For example, if the minimum distance values for the reference images R_k are calculated as {40, 60, 25}, the minimum distance values corrected by using the correction coefficients shown in Table 1 are {16, 30, 25}, so the reference image R_1 is selected as a reference image having the smallest minimum distance value.

In this embodiment, three-dimensional data of an object is measured, and this three-dimensional data is compared and compared with a reference image while the pose and illumination condition are corrected. Therefore, comparison and search can be performed at

high accuracy even when reference images of objects are captured under different conditions such as pose and illumination. Also, since three-dimensional data of an object is measured and this three-dimensional data is compared and compared with a reference image, comparison and search can be performed at high accuracy even when no three-dimensional object model of each object can be obtained beforehand, or even when only one or few reference images exist. Furthermore, since comparison is performed by correcting the comparison score fluctuations caused by the image sensing conditions of reference images, comparison and search can be performed at high accuracy even when the image sensing conditions of reference images are different or a reference image having poor image quality exists.

Note that in the second embodiment, the correction coefficient A_k is stored, and the distance value is corrected as it is multiplied by the correction coefficient A_k . However, the present invention is not limited to this embodiment. For example, it is also possible to store all the minimum distance values D_k^h between the representative three-dimensional models C_h and the reference images R_k , or store function parameters of an assumed distribution function. When a normal distribution is assumed, for example, it is possible to store an average value E_k and a standard deviation σ_k , and perform correction by $D_k' = (D_k -$

$E_k)/\sigma_k$ so that the distribution is normalized to a standard normal distribution.

(Third Embodiment)

Fig. 10 is a block diagram showing the third embodiment of the present invention. The same reference numerals as in Fig. 1 denote the same parts in Fig. 10. This embodiment comprises a three-dimensional data input means 10, reference image storing unit 30, pose estimating/comparing means 52, three-dimensional reference point extracting means 12, standard three-dimensional reference point storing unit 72, standard three-dimensional weighting coefficient storing unit 75, and reference weighting coefficient storing unit 77. The pose estimating/comparing means 52 includes a pose candidate deciding means 20, comparison image generating means 40, image comparing means 56, and input weighting coefficient converting means 70.

Outlines of the operations of these means are as follows. First, the three-dimensional data input means 10, reference image storing unit 30, pose candidate deciding means 20, and comparison image generating means 40 perform the same processes as in the first embodiment shown in Fig. 1.

Standard three-dimensional reference points corresponding to standard three-dimensional object models are stored in the standard three-dimensional reference point storing unit 72. Standard

three-dimensional weighting coefficients are stored in the standard three-dimensional weighting coefficient storing unit 75. Weighting coefficients corresponding to reference images are stored in the reference weighting coefficient storing unit 77.

The three-dimensional reference point extracting unit 12 manually or automatically extracts a three-dimensional reference point from three-dimensional data obtained from the three-dimensional data input means 10.

The pose estimating/comparing means 52 calculates a minimum distance value (or a maximum similarity degree) between the three-dimensional data obtained from the three-dimensional data input means 10 and a reference image obtained from the reference image storing unit 30, by using a weighting coefficient corresponding to the input data obtained from the input weighting coefficient converting means 70, or a weighting coefficient corresponding to the reference image obtained from the reference weighting coefficient storing unit 77.

More specifically, the pose candidate deciding means 20 generates at least one pose candidate. The comparison image generating means 40 generates a comparison image close to a reference image while projecting three-dimensional data onto a two-dimensional image in accordance with the pose candidate. The input

weighting coefficient converting means 70 obtains the coordinate correspondence of the standard three-dimensional weighting coefficient obtained from the standard three-dimensional weighting coefficient storing unit 75 to the three-dimensional data obtained from the three-dimensional data input means 10 by using the standard three-dimensional reference point obtained from the standard three-dimensional reference point storing unit 72 and the three-dimensional reference point of the three-dimensional data obtained from the three-dimensional reference point extracting means 12, and converts the standard three-dimensional weighting coefficient into a two-dimensional weighting coefficient in accordance with the pose candidate obtained from the pose candidate deciding means 20.

The image comparing means 56 includes a calculator 56a, selector 56b, and comparator 56c shown in Fig. 11. The image comparing means 56 calculates a distance value between the comparison image and the reference image in the calculator 56a by using the weighting coefficient obtained from the input weighting coefficient converting means 70 and corresponding to the input three-dimensional data or the weighting coefficient obtained from the reference weighting coefficient storing unit 77 and corresponding to the reference image, and selects a comparison image having

the smallest distance value for each reference image in the selector 56b, thereby estimating an optimum pose and obtaining a minimum distance value between the three-dimensional data and the reference image. In addition, in comparison (1 : 1 comparison) with one object (reference image), the comparator 56c compares the minimum distance with a threshold value to determine whether the two objects are the same. Also, in a process (1 : N comparison) of searching a plurality of objects for an object (reference image) closest to input three-dimensional data, the comparator 56c selects a reference image having the smallest minimum distance value.

The overall operation of 1 : N comparison of this embodiment will be explained in detail below with reference to Fig. 10 and a flowchart in Fig. 12.

Referring to Fig. 12, the three-dimensional data input means 10 inputs three-dimensional data (step 100). Then, the three-dimensional reference point extracting means 12 manually or automatically extracts a three-dimensional reference point from the three-dimensional data (step 170). The pose estimating/compared means 52 initially sets the image number of a reference image to $k = 1$ (step 150). The pose candidate deciding means 20 decides a pose candidate group $\{e_j\}$ (step 110).

Subsequently, the comparison image generating

means 40 generates a comparison image close to a reference image R_k while projecting the three-dimensional data onto a two-dimensional image in accordance with a pose candidate (step 120). The input
5 weighting coefficient converting means 70 obtains the coordinate correspondence of the standard three-dimensional weighting coefficient to the three-dimensional data by using the standard three-dimensional reference point and the
10 three-dimensional reference point of the three-dimensional data, and converts the standard three-dimensional weighting coefficient into a two-dimensional weighting coefficient in accordance with the pose candidate (step 180).
15 Then, the image comparing means 56 calculates a distance value between the comparison image and the reference image by using the weighting coefficient obtained from the input weighting coefficient converting means 70 and corresponding to the input
20 three-dimensional data, or the weighting coefficient obtained from the reference weighting coefficient storing unit 77 and corresponding to the reference image (step 131), and selects a comparison image having the smallest distance value for each reference image,
25 thereby estimating an optimum pose and obtaining a minimum distance value between the three-dimensional data and the reference image (step 140). Also, the pose

estimating/comparing means 52 increments the image number k by 1 (step 151), and compares the image number k with an image count M (step 152). If the image number k is equal to or smaller than the image count M , the
5 flow returns to step 110 to perform the same processing, thereby calculating the minimum distance value of the next reference image. Finally, if the image number k is equal to or larger than the image count M , a reference image R_k having the smallest minimum distance value is
10 output as a comparison result (step 153).

In 1 : 1 comparison, the processes in steps 100, 170, 110, 120, 180, 131, and 140 are performed in the same manner as in the first embodiment, and the process in step 155 of Fig. 3 is performed after that.
15 In step 155 of Fig. 3, whether the two objects are the same is determined by comparing the distance value with a threshold value as described earlier.

The operation of this embodiment will be explained in more detail below by using a practical
20 example. First, as in the explanation of the first embodiment, reference images $R_k(r)$ as shown in Fig. 5 are stored in the reference image storing unit 30. Also, standard three-dimensional reference points N_i^0 (i is the index of a reference point) corresponding to a
25 standard three-dimensional object model as shown in Fig. 13 are stored in the standard three-dimensional reference point storing unit 72. Three-dimensional

reference points are points for alignment. The example shown in Fig. 13 indicates five points, i.e., the middle point in the left eye, the middle point in the right eye, the tip of the nose, the left corner of the mouth, and the right corner of the mouth.

Although the three-dimensional reference points can be manually set in advance, it is also possible to automatically set them by using, e.g., a facial feature extracting method described in Marugame et al., "Extraction of Feature Portions from Facial Three-Dimensional Data Using Shape Information and Color Information", September, 2002, FIT (Forum of Information Science & Technology", 2002, I-100, pp. 199 & 200. The standard three-dimensional reference point can be calculated by the average coordinates of each three-dimensional reference point of a three-dimensional object model for learning prepared beforehand, or by a three-dimensional reference point obtained from a standard three-dimensional object model formed by averaging three-dimensional object models for learning.

Also, a standard three-dimensional weighting coefficient V_0^0 as shown in Fig. 14 is stored in the standard three-dimensional weighting coefficient storing unit 75. In this example shown in Fig. 14, a black region has a value of $V_0^0 = 1$, a white region has a value of $V_0^0 = 0$, and a hatched region has a value of $0 < V_0^0 < 1$. The standard three-dimensional weighting

coefficient can be calculated by using three-dimensional weighting coefficients of three-dimensional object models for learning prepared beforehand, i.e., by aligning the three-dimensional weighting coefficients such that the three-dimensional reference point of each three-dimensional object model for learning matches the standard three-dimensional reference point, and averaging the aligned three-dimensional weighting coefficients.

Points other than the reference points can be aligned by determining the correspondence of the reference points by interpolation or extrapolation, thereby setting transformation formulas $s_0 = Hs(s,t)$ and $t_0 = Ht(s,t)$ of coordinate values $\{s,t\}$ of a three-dimensional weighting coefficient and coordinate values $\{s_0,t_0\}$ of a standard three-dimensional weighting coefficient. A three-dimensional weighting coefficient of a three-dimensional object model for learning can be learned in advance by using images for learning formed by sensing the image of an object of the three-dimensional object model for learning under various conditions. More specifically, the image comparison system of the first embodiment is used to obtain an optimum pose by using a three-dimensional object model for learning as input three-dimensional data and an image for learning as a reference image, and calculate errors between pixels of the reference image

and a comparison image generated from the three-dimensional data when the optimum pose is obtained.

The weighting coefficient is an amount which represents the importance of a pixel in comparison. For example, the weight can be increased for a pixel having a small average error. The three-dimensional weighting coefficient can be set by calculating an average error by averaging errors between pixels of the comparison image and the reference image on the three-dimensional object model on the basis of the correspondence of the comparison image to the three-dimensional object model.

Also, weighting coefficients $U_k(r)$ as shown in Fig. 15 corresponding to reference images are stored in the reference weighting coefficient storing unit 77. In the examples shown in Fig. 15, a black region has a value of $U_k(r) = 1$, a white region has a value of $U_k(r) = 0$, and a hatched region has a value of $0 < U_k(r) < 1$. A weighting coefficient corresponding to a reference image is manually or automatically set in advance by, e.g., setting the weight of a region except for a facial region to 0, or decreasing the weight of a region where the luminance value is large or small.

Assume that three-dimensional data as shown in Fig. 6 is input from the three-dimensional data input means 10 (step 100 in Fig. 12). The three-dimensional reference point extracting means 12 manually or

automatically extracts a three-dimensional reference point from the three-dimensional data (step 170).

Fig. 16 shows examples of the three-dimensional reference point extracted from the three-dimensional data shown in Fig. 6.

Then, the pose estimating/comparing means 52 initially sets the image number of a reference image to $k = 1$ (step 150). The pose candidate deciding means 20 decides a pose candidate group $\{e_j\}$ (step 110). The comparison image generating means 40 generates a comparison image $G_{kj}(r)$ close to a reference image R_k while projecting the three-dimensional data onto a two-dimensional image in accordance with a pose candidate (step 120). Fig. 7 shows examples of comparison images generated for a reference image R_1 .

Subsequently, the input weighting coefficient converting means 70 obtains the coordinate correspondence of the standard three-dimensional weighting coefficient to the three-dimensional data by using the standard three-dimensional reference point and the three-dimensional reference point of the three-dimensional data, and converts the standard three-dimensional weighting coefficient V_0^0 into a two-dimensional weighting coefficient $W_{kj}(r)$ in accordance with the pose candidate (step 180). Fig. 17 shows examples of two-dimensional weighting coefficients generated for the comparison images shown in Fig. 7.

Then, the image comparing means 56 calculates a distance value D_{kj}' between the comparison image and the reference image by using the weighting coefficient $W_{kj}(r)$ obtained from the input weighting coefficient converting means 70 and corresponding to the input data, or the weighting coefficient $U_k(r)$ obtained from the reference weighting coefficient storing unit 77 and corresponding to the reference image (step 131), and selects a comparison image having the smallest distance value for each reference image, thereby estimating an optimum pose and obtaining a minimum distance value between the three-dimensional data and the reference image (step 140).

When a weighted Euclid distance is used, for example, the distance value D_{kj}' is calculated by

$$D_{kj}' = \sum_r W_{kj}(r) U_k(r) \{R_k(r) - G_{kj}(r)\}^2$$

It is also possible to use only one of the weighting coefficients $W_{kj}(r)$ and $U_k(r)$. Then, the pose estimating/compared means 52 increments the image number k by 1 (step 151), and compares the image number k with the image count M (step 152). If the image number k is equal to or smaller than the image count M , the flow returns to step 110 to perform the same processing, thereby calculating the minimum distance value of the next reference image. Finally, if the image number k is equal to or larger than the image count M in step 152, a reference image R_k having the

smallest minimum distance value is output as a comparison result (step 153).

In this embodiment, three-dimensional data of an object is measured, and this three-dimensional data is compared and compared with a reference image while the pose and illumination condition are corrected. Therefore, comparison and search can be performed at high accuracy even when reference images of objects are captured under different conditions such as pose and illumination. Also, since three-dimensional data of an object is measured and this three-dimensional data is compared and compared with a reference image, comparison and search can be performed at high accuracy even when no three-dimensional object model of each object can be obtained beforehand, or even when only one or few reference images exist. The accuracy of comparison and search can be further increased because image comparison is performed by weighted comparison using a weighting coefficient corresponding to each portion.

Note that although the third embodiment uses one standard three-dimensional weighting coefficient (and one standard three-dimensional reference point), a plurality of weighting coefficients may also be used. In this case, information indicating the correspondence of standard three-dimensional weighting coefficients to reference images is prestored. Also, the standard three-dimensional weighting coefficient is calculated by

an error average between pixels of an image for learning and a comparison image generated from a three-dimensional object model for learning. However, the present invention is not limited to this embodiment. Furthermore, in pose estimation, a weighted distance is calculated by using a weighting coefficient. However, it is also possible to calculate a distance without using any weighting coefficient in pose estimation, and calculate a weighted distance after an optimum pose is obtained.

(Fourth Embodiment)

Fig. 18 is a block diagram showing the fourth embodiment of the present invention. The same reference numerals as in Fig. 1 denote the same parts in Fig. 18. This embodiment comprises a three-dimensional data input means 10, reference image storing unit 30, pose estimating/comparing means 53, representative three-dimensional object model storing unit 36, three-dimensional comparing means 80, group storing unit 85, and reference image selecting means 82. The pose estimating/comparing means 53 includes a pose candidate deciding means 20, comparison image generating means 40, and image comparing means 55.

Outlines of the operations of these means are as follows. First, the three-dimensional data input means 10, reference image storing unit 30, pose candidate deciding means 20, comparison image generating

means 40, and image comparing means 55 perform the same processes as in the first embodiment shown in Fig. 1.

Also, representative three-dimensional object models prepared beforehand are stored in the
5 representative three-dimensional object model storing unit 36. Related information of representative three-dimensional object models and reference images (information indicating the correspondence of the representative three-dimensional object models to the
10 reference images) is prestored in the group storing unit 85.

The three-dimensional comparing means 80 compares three-dimensional data obtained from the three-dimensional data input means 10 with each
15 representative three-dimensional object model obtained from the representative three-dimensional object model storing unit 36, and selects the most similar representative three-dimensional object model. The reference image selecting means 82 selects a reference
20 image group corresponding to the representative three-dimensional object model selected by the three-dimensional comparing means 80, on the basis of the related information obtained from the group storing unit 85.

25 The pose estimating/comparing means 53 calculates a minimum distance value (or a maximum similarity degree) between the three-dimensional data

obtained from the three-dimensional data input means 10 and each reference image obtained from the reference image storing unit 30, and selects a reference image having the smallest minimum distance value. The

5 reference images herein used are those of the reference image group obtained from the reference image selecting means 82.

The overall operation of this embodiment will be explained in detail below with reference to Fig. 8 and flowcharts in Figs. 19 and 20. First, the
10 three-dimensional data input means 10 inputs three-dimensional data (step 100 in Fig. 19). Then, the three-dimensional comparing means 80 initially sets the model number of a representative three-dimensional
15 object model to $h = 1$ (step 210). The three-dimensional comparing means 80 calculates a similarity degree S_h between the three-dimensional data and each representative three-dimensional object model C_h (step 220). Subsequently, the three-dimensional comparing
20 means 80 increments the model number h by 1 (step 211), and compares the model number h with a model count H (step 212). If the model number h is equal to or smaller than the model count H , the flow returns to step 210 to perform the same processing, thereby calculating
25 the similarity degree of the next representative three-dimensional object model.

If comparison with all the representative

three-dimensional object models is completed in step 212, a model C_h having the largest similarity degree is selected (step 221). Then, the reference image selecting means 82 selects a reference image group
5 corresponding to the selected representative three-dimensional object model, on the basis of the related information obtained from the group storing unit 85 (step 230). Note that step 230 in Fig. 19 continues to step 150 in Fig. 20.

10 Subsequently, the pose estimating/comparing means 53 initially sets the image number of a reference image to $k = 1$ (step 150 in Fig. 20). The pose estimating/comparing means 53 determines whether a reference image R_k is included in a selected reference
15 image group L (step 240). If the reference image R_k is included in the reference image group, the flow advances to step 110; if not, the flow advances to step 151.

If the flow advances to step 110, the pose candidate deciding means 20, comparison image generating
20 means 40, and image comparing means 55 estimate an optimum pose and calculate a minimum distance value between the three-dimensional data and the reference image R_k by the same processes as in the first embodiment (steps 110 to 140). Then, the pose
25 estimating/comparing means 53 increments the image number k by 1 (step 151), and compares the image number k with an image count M (step 152). If the image number

k is equal to or smaller than the image count M, the flow returns to step 240 to perform the same processing. Finally, if the image number k is equal to or larger than the image count M, a reference image R_k having the smallest minimum distance value is output as a comparison result (step 153).

The operation of this embodiment will be explained in more detail below by using a practical example. First, as in the explanation of the first embodiment, reference images $R_k(r)$ as shown in Fig. 5 are stored in the reference image storing unit 30. Also, representative three-dimensional object models C_h as shown in Fig. 21 are stored in the representative three-dimensional object model storing unit 36. As shown in Table 2, the image numbers of upper candidates (reference images) obtained by comparing reference images with representative three-dimensional object models are stored in the group storing unit 85 in relation to these three-dimensional object models. That is, if the comparison results when the representative three-dimensional object models C_h are input to the image comparison system of the first embodiment are as shown in Table 3, for example, a list as shown in Table 2 is obtained by leaving reference image candidates having distance values of 40 or less behind.

[Table 2]

Model number	Reference image number
--------------	------------------------

1	2
2	1, 3

[Table 3]

Model number	Reference image number : distance value		
1	2 : 20	1 : 50	3 : 60
2	1 : 25	3 : 30	2 : 70

Assume that three-dimensional data as shown in Fig. 6 is input from the three-dimensional data input means 10 (step 100 in Fig. 19). The three-dimensional comparing means 80 initially sets the model number of a representative three-dimensional object model to $h = 1$ (step 210). Then, the three-dimensional comparing means 80 calculates the similarity degree S_h between the three-dimensional data and each representative three-dimensional object model C_h (step 220). As a means for comparing these three-dimensional data, it is possible to use an existing technique such as a three-dimensional shape identification apparatus described in Japanese Patent Laid-Open No. 4-119475.

Subsequently, the model number h is incremented by 1 (step 211), and the model number h is compared with the model count H (step 212). If the model number h is equal to or smaller than the model count H , the flow returns to step 210 to perform the same processing, thereby calculating the similarity

degree of the next representative three-dimensional object model. If comparison with all the representative three-dimensional object models is completed in step 212, a model C_h having the largest similarity degree is selected (step 221).

For example, if the similarity degrees of models C_h are $\{0.7, 0.9\}$, a model C_2 is selected. Then, the reference image selecting means 82 selects a reference image group $\{R_1, R_2\}$ corresponding to the selected three-dimensional object model C_2 from the list shown in Table 2 obtained from the group storing unit 85 (step 230). After that, the flow advances to processing shown in Fig. 20.

The pose estimating/comparing means 53 initially sets the image number of a reference image to $k = 1$ (step 150 in Fig. 20). The pose estimating/comparing means 53 determines whether a reference image R_k is included in the selected reference image group L (step 240). If the reference image R_k is included in the selected reference image group, the flow advances to step 110; if not, the flow advances to step 151.

In this example, the flow advances to step 110 if the reference image is R_1 or R_3 , and advances to step 151 if the reference image is R_2 . If the flow advances to step 110, the pose candidate deciding means 20, comparison image generating means 40, and image

comparing means 55 estimate an optimum pose and
calculate a minimum distance value between the
three-dimensional data and the reference image R_k by the
same processes as in the first embodiment (steps 110 to
5 140). Then, the pose estimating/comparing means 53
increments the image number k by 1 (step 151), and
compares the image number k with the image count M (step
152). If the image number k is equal to or smaller than
the image count M , the flow returns to step 240 to
10 perform the same processing. In this example, a minimum
distance value is calculated if the reference image is
 R_1 or R_3 . Finally, if the image number k is equal to or
larger than the image count M in step 152, a reference
image R_k having the smallest minimum distance value is
15 output as a comparison result (step 153).

In this embodiment, three-dimensional data of
an object is measured, and this three-dimensional data
is compared and compared with a reference image while
the pose and illumination condition are corrected.
20 Therefore, comparison and search can be performed at
high accuracy even when reference images of objects are
captured under different conditions such as pose and
illumination. Also, since three-dimensional data of an
object is measured and this three-dimensional data is
25 compared and compared with a reference image, comparison
and search can be performed at high accuracy even when
no three-dimensional object model of each object can be

obtained beforehand, or even when only one or few
reference images exist. Furthermore, high-speed search
can be performed because a reference image is selected
by comparison with a representative three-dimensional
5 object model.

Note that although one representative
three-dimensional object model is selected in the fourth
embodiment, a plurality of representative
three-dimensional object models may also be selected.
10 In this case, a union of reference image groups
corresponding to the individual representative
three-dimensional object models is a reference image
group.

(Fifth Embodiment)

15 Fig. 22 is a block diagram showing the fifth
embodiment of the present invention. The same reference
numerals as in Figs. 1 and 18 denote the same parts in
Fig. 22. This embodiment comprises a three-dimensional
data input means 10, reference image storing unit 30,
20 pose estimating/comparing means 53, representative
storing unit 31, second pose estimating/comparing means
(representative image selecting means) 54, group storing
unit 86, and reference image selecting means 82. The
pose estimating/comparing means 53 include a pose
25 candidate deciding means 20, comparison image generating
means 40, and image comparing means 55.

Outlines of the operations of these means are

as follows. First, the three-dimensional data input means 10, reference image storing unit 30, pose candidate deciding means 20, comparison image generating means 40, and image comparing means 55 perform the same
5 processes as in the first embodiment shown in Fig. 1.

Also, representative images prepared beforehand are stored in the representative image storing unit 31. A representative image can be a part of a reference image in the reference image storing unit
10 30, or a new image generated by, e.g., the average of reference images. When a representative image is a part of a reference image in the reference image storing unit 30, it is also possible to store only an image number and refer to the reference image in the reference image
15 storing unit 30.

Related information of representative images and reference images (information indicating the correspondence of the representative images to the reference images) is prestored in the group storing unit
20 86. The second pose estimating/comparing means 54 compares three-dimensional data obtained from the three-dimensional data input means 10 with each representative image obtained from the representative image storing unit 31, and selects the most similar
25 representative image. The reference image selecting means 82 selects a reference image group corresponding to the representative image selected by the second pose

estimating/comparing means 54, on the basis of the related information obtained from the group storing unit 86.

The pose estimating/comparing means 53
5 calculates a minimum distance value (or a maximum similarity degree) between the three-dimensional data obtained from the three-dimensional data input means 10 and each reference image obtained from the reference image storing unit 30, and selects a reference image
10 having the smallest minimum distance value. The reference images herein used are those of the reference image group obtained from the reference image selecting means 82.

The overall operation of this embodiment will
15 be explained in detail below with reference to Fig. 22 and flowcharts in Figs. 23 and 20. First, as shown in Fig. 23, the three-dimensional data input means 10 inputs three-dimensional data (step 100 in Fig. 23). Then, the second pose estimating/comparing means 54
20 initially sets the image number of a representative image to $h = 1$ (step 215). The second pose estimating/comparing means 54 calculates a similarity degree S_h between the three-dimensional data and each representative image R'_h (step 225).

25 Subsequently, the image number h is incremented by 1 (step 211), and the image number h is compared with an image count H (step 217). If the image

number h is equal to or smaller than the image count H , the flow returns to step 225 to perform the same processing, thereby calculating the similarity degree of the next representative image. If comparison with all
5 the representative images is completed in step 217, a representative image R'_h having the largest similarity degree is selected (step 226). Then, the reference image selecting means 82 selects a reference image group corresponding to the selected representative image, on
10 the basis of the related information obtained from the group storing unit 86 (step 235). Note that step 235 in Fig. 23 continues to step 150 in Fig. 20.

The pose estimating/comparing means 53 initially sets the image number of a reference image to
15 $k = 1$ (step 150 in Fig. 20). The pose estimating/comparing means 53 then determines whether a reference image R_k is included in the selected reference image group (step 240). If the reference image R_k is included in the selected reference image group, the flow
20 advances to step 110; if not, the flow advances to step 151.

If the flow advances to step 110, the pose candidate deciding means 20, comparison image generating means 40, and image comparing means 55 estimate an
25 optimum pose and calculate a minimum distance value between the three-dimensional data and the reference image R_k by the same processes as in the first

embodiment (steps 110 to 140). Subsequently, the pose
estimating/comparing means 53 increments the image
number k by 1 (step 151), and compares the image number
 k with the image count M (step 152). If the image
5 number k is equal to or smaller than the image count M ,
the flow returns to step 240 to perform the same
processing, thereby calculating the minimum distance
value of the next reference image. Finally, if the
image number k is equal to or larger than the image
10 count M in step 152, a reference image R_k having the
smallest minimum distance value is output as a
comparison result (step 153).

The operation of this embodiment will be
explained in more detail below by using a practical
15 example. First, as in the explanation of the first
embodiment, reference images $R_k(r)$ as shown in Fig. 5
are stored in the reference image storing unit 30.
Also, representative images R'_h as shown in Fig. 24 are
stored in the representative image storing unit 31. As
20 shown in Table 4, the image numbers of upper candidates
(reference images) obtained by comparing reference
images with representative images are stored in the
group storing unit 86 in relation to these
representative images. In this comparison, the existing
25 image comparison system described in, e.g., patent
reference 1 or 2 described previously can be used.
[Table 4]

Image number	Reference image number
1	2
2	1, 3

Assume that three-dimensional data as shown in Fig. 6 is input from the three-dimensional data input means 10 (step 100 in Fig. 23). The second pose estimating/compared means 54 initially sets the image number of a representative image to $h = 1$ (step 215). Then, the second pose estimating/compared means 54 calculates the similarity degree S_h between the three-dimensional data and each representative image R'_h (step 225). Subsequently, the second pose estimating/compared means 54 increments the image number h by 1 (step 211), and compares the image number h with the image count H (step 217). If the image number h is equal to or smaller than the image count H , the flow returns to step 215 to perform the same processing, thereby calculating the similarity degree of the next representative image.

If comparison with all the representative images is completed in step 217, a representative image R'_h having the largest similarity degree is selected (step 226). For example, if the similarity degrees of representative images R'_h are $\{0.7, 0.9\}$, a representative image R'_2 is selected. Then, the reference image selecting means 82 selects a reference image group $\{R_1, R_3\}$ corresponding to the selected

representative image R'_2 from the list shown in Table 4 obtained from the group storing unit 86 (step 235). After that, the processing shown in Fig. 20 is performed.

5 The pose estimating/comparing means 53 initially sets the image number of a reference image to $k = 1$ (step 150 in Fig. 20). The pose estimating/comparing means 53 then determines whether a reference image R_k is included in the selected reference
10 image group (step 240). If the reference image R_k is included in the selected reference image group, the flow advances to step 110; if not, the flow advances to step 151.

 In this example, the flow advances to step 110
15 if the reference image is R_1 or R_3 , and advances to step 151 if the reference image is R_2 . If the flow advances to step 110, the pose candidate deciding means 20, comparison image generating means 40, and image
20 comparing means 55 estimate an optimum pose and calculate a minimum distance value between the three-dimensional data and the reference image R_k by the same processes as in the first embodiment (steps 110 to 140). Subsequently, the pose estimating/comparing means 53 increments the image number k by 1 (step 151), and
25 compares the image number k with the image count M (step 152). If the image number k is equal to or smaller than the image count M , the flow returns to step 240 to

perform the same processing. In this example, a minimum distance value is calculated if the reference image is R_1 or R_3 . Finally, if the image number k is equal to or larger than the image count M in step 152, a reference
5 image R_k having the smallest minimum distance value is output as a comparison result (step 153).

Note that although one representative image is selected in the fifth embodiment, a plurality of representative images may also be selected. In this
10 case, a union of reference image groups corresponding to the individual representative images is a reference image group.

In this embodiment, three-dimensional data of an object is measured, and this three-dimensional data
15 is compared and compared with a reference image while the pose and illumination condition are corrected. Therefore, comparison and search can be performed at high accuracy even when reference images of object are captured under different conditions such as pose and
20 illumination. Also, since three-dimensional data of an object is measured and this three-dimensional data is compared and compared with a reference image, comparison and search can be performed at high accuracy even when no three-dimensional object model of each object can be
25 obtained beforehand, or even when only one or few reference images exist. Furthermore, high-speed search can be performed because a reference image is selected

by comparison with a representative image.

Note that in each of the first to fifth
embodiments, three-dimensional data (model) has as its
information a shape and texture in a three-dimensional
5 space (x,y,z) on the object surface. However, the
present invention is not limited to these embodiments as
long as equivalent information is obtained. For
example, it is also possible to use a distance image
which expresses a distance to the object surface in a
10 certain direction as an image, and a texture image
captured in the same direction.

The functions of the individual means as
constituent elements of the image comparison system of
the present invention can of course be implemented as
15 hardware, and can also be implemented by loading an
image comparison program (application) which executes
the functions of the means of the first to fifth
embodiments into a memory of a computer processor,
thereby controlling the computer processor. This image
20 comparison program is stored in a recording medium such
as a magnetic disk or semiconductor memory, and
implements the functions described above by controlling
the operation of a computer processor when loaded from
the recording medium into the computer processor.

25 Industrial Applicability

The present invention is applicable to
applications such as an image comparison system which

searches for an image of an object such as a human face from a database, and a program which implements the image comparison system on a computer.

The present invention is also applicable to
5 applications such as search of an image of an object such as a human face existing on a network or on the Internet. Furthermore, the present invention can be suitably applied to applications such as determination of whether an image such as an identification photograph
10 and a person holding it are the same.